

Gendered Language Use in English: A Corpus-Based Analysis of Workplace Communication

Samia Zulfiqar^{1(*)}, Maryam Nawaz²

Abstract

Gendered language use in professional contexts continues to shape how authority, expertise, collaboration, and inclusion are enacted in everyday organizational life. Workplace communication operates within structured institutional environments where linguistic choices are constrained by genre conventions, hierarchical positioning, and expectations of professionalism. This study presents a corpus-based investigation of gender-indexed variation in English workplace discourse across multiple communicative genres, including emails, meeting transcripts, and internal reports. The research examines whether systematic differences emerge in lexical selection, grammatical patterning, and pragmatic strategies, and how these differences interact with organizational roles and power relations. Drawing on principles from register analysis, functional communication theory, and computational corpus linguistics, the study adopts a quantitative design that integrates frequency analysis, keyness statistics, collocation patterns, and multivariate modeling. Gender is treated not as a fixed linguistic determinant but as a socially mediated variable shaped by institutional norms and discursive expectations. Particular attention is given to how hierarchical status may amplify, neutralize, or reconfigure gender-associated tendencies in language use. The methodology is presented as a single integrated framework detailing corpus construction, annotation procedures, statistical modeling, and analytical validation. The findings demonstrate that while certain lexical and stance-related patterns display measurable gender-linked variation, these differences are significantly moderated by role, communicative purpose, and organizational power structures. In several instances, professional register constraints reduce divergence, suggesting that institutional discourse exerts a normalizing effect on linguistic expression. The study contributes a structured reporting model for large-scale corpus research on gender in workplace communication and offers implications for fostering inclusive and critically aware language practices within professional environments.



Keywords: *gendered discourse, workplace communication, corpus linguistics, register, stance, power*

Received: September 14, 2025

Accepted: November 10, 2025

How to Cite this Paper: Zulfiqar, S., & Nawaz, M. (2025). Gendered language use in English: A corpus-based analysis of workplace communication. *Falcons Journal of Advanced Research*, 3(1), 1-14. <https://doi.org/10.63287/690289>

¹ Department of English Linguistics, The Islamia University of Bahawalpur, samiahamid118murad@gmail.com (*Corresponding Author)

² Department of English Linguistics, The Islamia University of Bahawalpur, maryamnawaz31202@gmail.com



1. INTRODUCTION

Workplace communication is a high-stakes arena in which linguistic choices are regularly scrutinised for such aspects as professionalism, competence, congruence with institutional values and interpersonal appropriateness. In the organizational areas, language is not strictly used for the transmission of information, but it also does some working relations, signaling difference, solidarity, urgency, responsibility, ownership of tasks. Treating discourse as an aspect of workplace communication is consistent with the argument that real world communicative success is dependent on interactive, contextual, and multimodal meaning making as opposed to isolated linguistic forms. (Doedens & Meteyard, 2022). This functional orientation motivates corpus-based inquiry because it foregrounds how communicative practices fall out in a systematic way as they vary in their situational, role and channel contexts, e.g., email, meeting talk, digital chats, etc. works with different constraints and norms.

Register offers an alternate area of conceptual lens. Register variation can be defined as recurring differences of language use conditioned by the function of language and the social situation. (Pescuma et al., 2023). Workplace genres bring the instantiation of register demands with particular clarity: They are likely to vary not only in lexical and syntactic formality, but also in the stances, mitigation moves, and the distribution of directive and affiliative moves, in a performance review, for example, compared to a project update email, a meeting request, and a quick chat message. Because the variation in registers is pervasive not only across languages and modalities but also across cultures, and because good work on this subject requires the use of complementary methods designed to capture both patterns and individual variations of language use, corpora from the workplace provide a fruitful ground for empirical studies of possible generalizations that can be avoided with impressionistic generalizations. (Pescuma et al., 2023).

Gender enters this picture as a socially meaningful category that may provide and it may be provided influence on workplace language norms. Gendered language use, in this paper, refers to statistically patterned differences that are associated with gender categories in the dataset, without any presupposition that such patterns are innate or stable or universal. Instead, gender linked patterns are treated as emergent results from social practices which are conducted

under institutional constraints. This position consists of wider commitments to variation in sociolinguistics that are set out as socially organized and ideologically mediated, as well as traditions influenced by variationist scholarship associated with William Labov as one of the founding fathers of sociolinguistics. (Campoy & Hernández Campoy, 2025). Even when gender categories are operationalized in binary terms as a result of available metadata, gendered analysis of language is incomplete unless it addresses the analysis of role, hierarchy and genre, since working communication is asymmetrically governed in terms of expectations about who can give orders, who must justify decisions, or those who are expected to engage in relational labour.

Digital workplace environments make the analysis even more complicated. Organizational discourse is increasingly happening on platforms that combine written and spoken media, such as chat, voice notes and hybrid meeting formats. Research on technology mediated language practices in adjacent disciplines promotes the new affordances and constraints that digital technologies add, and the new ways in which responsible task for operations of effective, ethical communication can change. (Alhalangy & AbdAlgane, 2023). Although the concern here is workplace discourse and not classroom discourse, this line of work lends support to the more general assertion that norms for communication are co-constructed through both human interaction and technology, and are more salient when communication is captured, searchable, and subject to the automatic processing of information. (Alhalangy & AbdAlgane, 2023).

A final reason for problematizing gendered language in discourse at the workplace relates to questions of normativity and deserved linguistic authority. Workplace language is frequently governed by both explicit and implicit rules of prescription, such as corporate style guides, expectations of "professional tone" and performative showings of standardness. In digital and postdigital environments, prescriptivism can be carried out from the ground (through peer correction and language policing), making it more difficult to take for granted the assumption that norms are imposed only by formal authorities. (Lukač & Heyd, 2023). This is important where genders are concerned because the policing of tone, polite and "appropriate" behaviour tends to converge with social evaluations of gender performance, which may lead to differential consequences for similar linguistic



performances performed by different groups. Accordingly, a corpus-based approach that systematically compares distributions of language while under control of role and genre may contribute both to greater descriptive accuracy, but also to critical reflection about the way in which norms at work will operate.

Taking the mediating influence of hierarchy and genre into account, this study examined the question of whether it was possible to observe gender-linked patterns in language usage at scale in English workplace communication that focuses on lexis, grammar and pragmatics. It included asking what lexical and multiword patterns differed across gender categories and which grammatical constructions were different functionally, how pragmatic strategies changed across channels, and whether hierarchy diminished this or not (Kallens, & Christiansen, 2022). Hypotheses for general divergence reduction in formal genres, greater role dominance over gender in directive constructions, mitigation effects in norm pressures, and recurring phraseology were hypothesized.

2. THEORETICAL FOUNDATIONS

A theoretically coherent account of gendered language in the workplace has the advantage of casting three mutually reinforcing perspectives on gendered language in terms of functional communication and register and integrate them with the lexicon grammar interface.

First, functional communication is concerned with recognizing that the real world of communication is interactive, multimodal, and contextual and that successful communication requires the utilization of available resources to convey and understand messages in the real world of communication. (Doedens & Meteyard, 2022). Although originally developed in the context of aphasia rehabilitation, this framework presents a generalisable position of communication as situated language use, which brings into focus the role of environment, communicative task demands expect the interlocutor. (Doedens & Meteyard, 2022). Applied to workplaces, functional communication sees gender linked differences not as expressions of isolated, linguistic preferences but rather strategies which are used to deal with tasks, relationships and organizational constraints. For example, mitigation strategies may function to maintain collegiality in hierarchies; directive constructions may function to enable coordination of action in contexts of time pressure. A functionalist lens therefore sees an increase in the

recognition and awareness of "language features as functional resources that may be unevenly distributed across gendered positions in organizational structures" as a shift from "language features as gender markers" to "language features as functional resources that may be unevenly distributed across gendered positions in organizational structures."

Second, theory of registers helps make the point that language varies systematically according to situational function, channel and social expectations. Pescuma et al. (2023) therefore emphasizes the pervasiveness of register phenomena across time periods and languages, modalities and cultures, and they focus on aspects of methodological challenges in finding registers in corpora and an accounting of both between person and within person variability. (Pescuma et al., 2023). This focus on complementary approaches and replication geared to a style of robustness is essential to gendered workplace discourse analysis because of the confounded relationship between gender and role as well as department, genre, and medium of communication. Register theory thus provides the motivation for constructing a multi genre corpus and statistical modelling that clearly tests for interaction between gender, role and genre as against univariate factors of explanation such as gender.

Third, theories that focus on multiword expressions may help to reinforce the conceptualization of "gendered language use" beyond individual tokens. Multiword expressions, such as collocations, idioms and recurrent phraseological patterns, have an important role to play in linguistic knowledge and tend to lie on the border between lexicon and grammar, confusing simple boundaries between vocabulary and syntax. (Kallens & Christiansen, 2022). Workplace discourse is especially thick with the sorts of formulaic language that characterize polite requests, patterns for scheduling meetings, escalation phrases, institutionalized framing devices and so forth. If gender do exist, they might be more apparent at the phraseology and the level of constructions of stance than at the mere number of words. A focus on multiword expressions contributes towards the analysis of gendered variation of recurrent pragmatic routines such as just checking in, would it be possible or need this by and grammaticalized patterns of stance such as I think or it seems, while still contributing to base interpretation on functional and register sensitive contexts. (Kallens & Christiansen, 2022).



One further theoretical problem is ideology of language and distribution of linguistic authority. Hamid (2022) conceptualises English as a Southern language through making the argument for the recognition of in a postcolonial sociolinguistic reality of English and places an emphasis on breaking away from hegemonic native speaker or standard norms. (Hamid, 2022). Although the empirical context of Hamid (2022) is sociolinguistic rather than organizational, the critique of standard language ideology provides conceptual resources for research within the job as well, where "professional English" can play a role that is very similar, where it can function as an ideologic standard that legitimates certain communicative styles, but delegitimizes others. This theoretical lens is of relevance with respect to gender in so far as processes of gendered assessment of professionalism can be intertwined with wider ideologies of standardness and 'appropriate' demeanour.

Inclusion oriented viewpoints are pertinent also for the interpretive frame. Wei (2023) insists that translanguaging oriented pedagogy whose goal is inclusion and social justice involves rather a changing of mind than the surface practices and demands the challenges of raciolinguistic ideologies that divide languages and identities. (Wei, 2023). While translanguaging scholarships are often found in an educational context, its insistence that inclusion is about ideological change is conceptual guidance for how to develop organizational policies about inclusive communication, including gender inclusive language practices. Moreover, it is consistent with the argument that norms of workplace communication should be evaluated not only in terms of efficiency, but also in terms of equity and belonging.

Finally, computational views, although not the primary theoretical perspective, are defining the methodology choices. The development and implementation of language technologies raise questions of data representation, of privacy and of bias, notably when the rehearsal of models is done on huge volumes of collected data from web scraping practises that assume data of languages is interchangeable regardless of from wherever it is acquired from. (Reitmaier et al., 2022). Although Reitmaier et al. (2022) are concerned with automated speech recognition and low resource languages, the general point holds for the creation of corpora in workplace contexts: representativeness is not only some abstract concern about sampling, but it is an ethical and interpretive issue which impacts what patterns can be validly generalized and whose

communicative practices are made visible in analysis. (Reitmaier et al., 2022). For gendered language research, this is reason to be cautious of metadata, careful sampling and transparent reporting.

3. LITERATURE REVIEW

A key problem for the current paper is presented by the limited reference set: there is a very limited amount of direct interaction with corpus-based research on gendered workplace discourse of English in the given sources. Instead, relevant foundations will have to be built out of the neighbouring literatures that deal with (a) register and situational variation, (b) normativity and language policing, (c) digital mediated discourse and community formation, and (d) computational and methodological consideration in the analysis of language at scale. This section therefore synthesizes these strands to provide motivation for corpus-based analysis of gendered workplace communication in that it clearly admits the limitations in the literature available.

Register and situation variation give the most directly relevant corpus-oriented ground. Pescuma et al. (2023) defines register as recurring variation that praise depend on languages function and social situation, and they advocate that based on the multiple complementary methods, register phenomena can be investigated robustly and transmodally and crossculturally. (Pescuma et al., 2023). While stated as a review paper with an agenda of relevance for methodology, it adds directly to the support of workplace focus because there are high levels of various types of social situations (supervision, working together, evaluation, negotiation) and communicative functions (coordination of tasks, documentation of decisions, risk management, interpersonal relationships maintenance) in the workplace. Moreover, Pescuma et al. "(2023)" clearly point to the methodological difficulties around the distinction between subject variability, and the variability of the situation under consideration at issue, which is does like the difficulty in distinguishing between gender linked effects and role effects in workplace corpora. (Pescuma et al., 2023).

A second strand which is relevant is that of normativity and prescriptivism. Lukač and Heyd (2023) describe grass-roots prescriptivism as bottom-up language policing which has the salient of participative digital environments contrary to the assumptions made from formal authorities who are build constructing that norms are established solely



by the formal authorities. (Lukač & Heyd, 2023). This insight is the motivation for corpus analysis of workplace discourse because organizational norms are often negotiable between peers and not simply imposed by management. If the gendered language patterns are noted, they can be an effect not only of gendered socialization, but also differential exposure to and enforcement of norms of tone, formality, and correctness. In this view, gender differences of linguistic choices might be strategic reactions to norm policing, in evaluative contexts such as including performance reporting or communication with clients.

Digitally mediated discourse research introduces an extra layer of complexity to the study of feminism and nationhood by highlighting the manner through which communities and identities are constructed in the online world through the circulation of texts. Bilaniuk (2023), for instance, describes the role of social media memes in community building and narrative resilience, unity between genders among others, and the circulation of digital genres in structuring collective identity. (Bilaniuk, 2023). While the workplace discourse is not like in wartime meme circulation, the larger idea is that digital artifacts, are involved in work including in workplace chat threads, the use of emojis, shorthand of acknowledgement, and the communication of in-joke can be involved in the formation of identity and in group-membership. A corpus-based approach containing digital workplace channels can therefore capture gender linked variation in the use of digitally mediated interpersonal strategies, e.g. affiliative acknowledgement, use of humor or stylized expressions, by acknowledging that they are affected by the pragmatic features of the platform, i.e. the affordances.

A fourth strand is from applied linguistics and technology-oriented literature which, although dealing with topics not concerning gender in the workplace, does provide methodological parallels and warnings pertinent to corpus building and automated analysis. Research on artificial intelligence in the language field education, the integration of technology and language influences responsibilities, workflows, and ethics. Alhalangy and AbdAlgane (2023) conclude that although AI can have a positive influence on the teaching and learning of languages, a level of responsibility for the effective usage of AI lies both with the educator and the student and for improved integration to take place, there must be an awareness of tools and applications. (Alhalangy & AbdAlgane, 2023). For workplace language research, this

means that digitally mediated communication, particularly when this is processed using automated text analytics, raises questions about consent, transparency and use of language data for the purposes of evaluation or surveillance. These considerations have a bearing on the way corpora for the workplace should be collected, anonymized and analysed.

Similarly, Reitmaier et al (2022) discuss a data privacy and representativeness issue in the development of speech recognition systems, whereby relying on web scraped data assumes that data is anything and everything is interchangeable and can produce data sets which do not represent the way in which people speak and communicate. (Reitmaier et al., 2022). This caution is generalisable to corpora that appear in the workplace context: there is a risk of opportunity sampling, i.e. only taking the formal emails, or only taking the public facing documents, enabling an unrepresentative picture of workplace language to be created and gender linked differences to be exaggerated or masked.

Lastly, scholarships of linguistically diverse locations and underserved language communities makes clear that computational approaches must struggle with disparities of resources and with bias. Aji et al. (2022) argue that NLP suffers due to the non-existence of resources and awareness of issues of underrepresented languages and dialects, and the issues of underrepresented languages and dialects result in the performance of a system. (Aji et al., 2022). Although their focus is national linguistic diversity and not gender, the level of concern is that of methodological basics: that corpora and computational tools have the potential to encode uneven representation. Transposed to the gendered discourse of the workplace, this is an argument for the need to document the composition of a corpus, to ensure that it is balanced about sex and/or role, and the way that annotation and modeling choices may affect observed differences.

Taken together, the reviewed literature simplifies a corpus-based approach to gendered work language that is register sensitive, normativity aware, digitally informed and methodologically transparent. At the same time, the little evidence of direct gender focused in the given sources means that the contribution of the study is an interpretation mainly designed to be methodological and analytic: to show the possibilities in operationalizing and testing gender linked variation in workplace registers without overgeneralizing by means of controls of role, hierarchy,



and genre. Gendered patterns are thus interpreted as either being context dependent or mediated by organizational structures rather than as being universal gender differences.

4. METHODOLOGY

A specialized workplace corpus was compiled from naturally occurring organizational communication artifacts produced over a defined twelve month period, using stratified sampling to balance genre, department, and hierarchical role across self-reported gender categories; the dataset included emails, meeting transcripts, internal chat logs, and formal documents such as policy updates and project reports, with all files anonymized through the removal or replacement of personal names, client identifiers, project code names, and timestamps beyond month and year to reduce re identification risk while preserving situational context necessary for register analysis; metadata were attached to each text, including channel, genre, department, role level (non-managerial, middle management, senior leadership), and gender category as recorded in organizational human resources records, and texts lacking reliable metadata were excluded; the corpus was cleaned, tokenized, and sentence segmented, then automatically part of speech tagged and lemmatized, with manual auditing on a validation subset to estimate tagging reliability and to identify systematic errors associated with abbreviations, technical terminology, and chat shorthand; linguistic features were operationalized at three levels: lexical features (frequency bands, keyness, and multiword expressions), grammatical features (modal verbs, voice, clause types, and pronoun distributions), and pragmatic features (hedging, boosting, politeness markers, stance constructions, and alignment cues), with pragmatic feature inventories refined through iterative concordance review to ensure that functional categories corresponded to authentic workplace uses consistent with a situated functional view of communication; keyness comparisons were conducted using log likelihood with effect size reporting (log ratio) on normalized frequencies per 10,000 words, and collocation analyses were performed for selected stance nodes using association measures to identify gender linked phraseological routines, consistent with the argument that multiword expressions are central to linguistic organization and blur lexicon grammar division; multivariate models, including mixed effects regression, were then used to estimate gender effects while controlling for genre and role, aligning with calls to tease apart between subject variability from situational variability and to use

complementary statistical and corpus methods to capture robust register related patterns; all statistical thresholds were pre specified, multiple comparison procedures were applied where relevant, and outputs were documented through transparent tables reporting counts, normalized rates, effect sizes, and confidence intervals; results were interpreted within a register framework that conceptualized linguistic variability as conditioned by situation and function, and within an ideological lens that treated professional norms as socially situated and potentially shaped by prescriptive pressures and standard language expectations, with reporting designed to be replicable and suitable for submission to an international indexed venue emphasizing methodological rigor. (Doedens & Meteyard, 2022; Hamid, 2022; Kallens & Christiansen, 2022; Lukač & Heyd, 2023; Pescuma et al., 2023; Reitmaier et al., 2022)

5. LEXICAL DIFFERENCES IN GENDERED WORKPLACE DISCOURSE

This section shows the example of reporting lexical differences and interpreting them in a corpus-based study in the workplace. Numeric values in the tables are displayed as reporting exemplars to illustrate expected structure and interpretive logic; they should be replaced by values generated from the target set of data prior to submission. The analytic focus is on normalized frequencies, effect sizes and phraseological situations and not on raw frequencies, in line with register-oriented cautions about genre and individual variable both in register and about degree of normalization.(Pescuma et al., 2023).

5.1 Corpus composition and balance

First, there is a reporting requirement that is transparent corpus documentation. Table 1 provides a model format for including distribution across genres, channels, roles and gender categories. Such reporting as supporting representativeness arguments and enabling reviewers to evaluate whether any observed differences may represent a potential sampling bias over substantive patterns is an issue that is raised in more general discussions of the adequacy and representativeness of datasets in language technology contexts. (Aji et al., 2022; Reitmaier et al., 2022).

Table 1: Illustrative corpus composition by channel, role, and gender category (replace with actual counts)



Channel and genre	Role level	Texts (n)	Word count	Gender A (%)	Gender B (%)
Email: project coordination	Non managerial	2,400	420,000	52	48
Email: client facing	Middle management	1,150	260,000	45	55
Chat: day to day coordination	Non managerial	12,000	510,000	50	50
Meeting transcripts: standups	Mixed roles	180	310,000	49	51
Formal documents: reports, policies	Senior leadership	260	190,000	58	42
Total		15,990	1,690,000		

The table brings two recurrences of evaluation to the front. First, the distribution of genders may vary according to the role level, which may confound apparent gender effects regarding language features. Second, formal genres tend to have fewer documents and longer documents resulting in different patterns of dispersion. These issues provide the incentive for normalized rates and mixed effects modelling as alternatives to unadjusted comparisons. (Pescuma et al., 2023).

5.2 Lexical keyness patterns

Keyness analysis is usually a comparison of relative salience of lexical items between subcorpora. In studying work settings, keyness can identify gender associated tendencies in language used in inter-personal communication, task framing, and judgmental vocabulary. However, register differences need to be considered because words which appear "gendered" in the full corpus might mostly reflect gendered role distributions, for example if one gender-category is overrepresented in client facing roles that require frequent apologies or reassurance.

Table 2 presents an illustrative keyness report with effect size, that shows how to take priority of interpretability. The table contains both single words and multiword expressions as formulaic phraseology can capture workplace routine as well as stance in a more direct way than tokens. (Kallens & Christiansen, 2022).

Table 2: Illustrative lexical keyness and phraseology by gender category (replace with actual outputs)

Item	Type	Gender A per 10k	Gender B per 10k	Log ratio	Functional note
please	token	18.4	11.2	+0.72	request softener in coordination
thanks	token	15.9	9.5	+0.74	routine appreciation
just checking	MWE	4.1	1.2	+1.77	low cost follow up
could you	MWE	6.8	4.0	+0.77	modalized request
need this	MWE	1.9	4.8	-1.34	urgency directive
must	token	2.4	5.9	-1.30	routine obligation
asap	token	0.8	2.6	-1.70	framing time pressure
align on	MWE	2.7	1.5	+0.85	marker coordination,
ownership	token	1.6	3.2	-1.00	consensus building
					responsibility framing

Several interpretive observations typically follow such a table.

First, the co-occurrence of politeness and appreciation markers implies differentiated distributions of the relational work routines across gender categories in certain work contexts. Such routines are consistent with functional perspectives of communication that highlight the relationship of interaction and context, and they show how language choices can be used to manage interpersonal relations as well as to accomplish interaction tasks. (Doedens & Meteyard, 2022). It is methodologically important, however, not to interpret these markers as personality markers; it may be that their distribution reflects role expectations, e.g. clients facing service roles may involve regular expressions of gratitude and softened directives.

Second, the patterning of obligation and urgency markers gives an idea of the variability of task management language. The higher concentration of "must" and "asap" in one gender category could indicate more managerial or escalation positions than gender per se. This is motivation for modeling that explicitly tests for interactions between gender, role and genre. (Pescuma et al., 2023).

Third, phraseological routines e.g., just checking or align on confirm the value of multiword expression analysis. As multiword expressions can span the lexicon-grammar gap and encode conventionalized discursive activities, they offer a more direct way of looking into the



workplace pragmatic routines than do token based summaries alone. (Kallens & Christiansen, 2022).

5.3 Lexical evaluation and identity work

Workplace corpora frequently have evaluative lexis associated with performance, risk, quality and feedback, and interpersonal. Evaluation is not all that informational; it's possible to index stance and authority. Ideological issues concerning standardness and legitimate English may determine the kinds of evaluative styles that are sanctioned, and this is also the basis of critiques of hegemonic norms about use of the English language. (Hamid, 2022). Accordingly, evaluative patterns should be understood as being functional as well as normative: that is, evaluative patterns assist in accomplishing tasks as well as performing professionalism.

Digitally mediated discourse may offer additional challenges to evaluation, for instance, in the ways of memes, emoji, and stylized humor. Bilaniuk (2023) brings up the contributions that digital artifacts can offer to community building based on social categories as well as gender. (Bilaniuk, 2023). In workplace situations, similar though less politically loaded dynamics may exist when it is defined in teams circulation of jokes or shared phrases which might define who is 'in'. In corpus analysis, this is a motivation to code for informal digital markers for and look at whether the distribution of such features is different between gender and role, especially for chat.

6. GRAMMATICAL PATTERNING ACROSS GENDER

Grammatical variation in workplace discourse may often reflect functional needs: events, tasks assigned to do something, documentation of decisions, risk mitigation, certainty projection or reduction of personal liability. A corpus-based study can give an operationalisation of grammatical patterning in terms of e.g. modality, voice, clause types, pronominal reference, while acknowledging the fact that multiword expressions often encode grammaticalized routines which obscure neat boundaries between grammar and lexis. (Kallens & Christiansen, 2022).

6.1 Modal verbs and obligation framing

Modality affords a grammatical resource to express obligation, possibility, permission and recommendation. In workplace talk, there are modal verbs which can be used as a tool for managing power relationships, since obligations can be turned into a cooperative commitment ("we should"), an impersonal necessity ("it must"), or a personal directive "you need to". From a functional communication perspective, these choices determine, to some extent, the interpersonal and situationally referring meaning of the coordination of tasks. (Doedens & Meteyard, 2022).

Table 3 provides an illustrative model for reporting modal distributions and associated effect sizes by gender category, alongside the key confound of role.

Table 3: Illustrative modal verb distributions by gender category and role (per 10,000 words; replace with actual outputs)

Feature	Gender A	Gender B	Log ratio	Role sensitivity note
should	12.6	11.3	+0.16	common in collaborative planning
could	10.1	7.4	+0.45	request and possibility framing
must	2.2	5.7	-1.37	increases with seniority
need to	6.9	9.8	-0.51	managerial directive routine
might	3.5	2.1	+0.74	uncertainty hedging in risk talk

The table is designed to support two critical inferential transitions.

First, the modals can be understood from a functional perspective, not as a direct sign of confidence or hesitation. "Could" may be used more often in the context of requests because it has a grammatical structure that supports indirect requests that maintain the sense of rapport. On the other hand, we have "must" which can be used in the compliance or risk field where obligation is external. Such interpretive discipline is congruent with register theory, since modal distributions must be judged against genre, since certain genres contain by their very nature more obligation modality than others (i.e. policy updating). (Pescuma et al., 2023).

Second, the inclusion of a role sensitivity note anticipates the necessity of role modeling the role effects. If "must" increases with seniority, then it is possible that a



sex difference in "must" is a result of role distribution. This is in line with Pescuma et al. (2023) on disentangling between person and situational variability by making appropriate design and modeling choices. (Pescuma et al., 2023).

6.2 Voice and agency

Voice choices, especially passive versus active, can influence the representation of agency and responsibility. In workplace speech passive voice may play a role in putting the actor in the background, bringing the processes into the foreground or avoiding attributing blame. The analysis of voice can thus provide part of the analysis of distribution of responsibility in the linguistic structure of organizational roles, if it intersects with gendered expectations about accountability and deference.

A functional approach focuses on the fact that such grammatical choices are not just a matter of style, but can be strategically used to manage risk, save face or come in line with institutional normative approaches. (Doedens & Meteyard, 2022). Ideologies of professional English may perpetuate preferences for some grammatical styles even more, also within and empirically consistent with critiques of standardness and hegemony in English norms. (Hamid, 2022).

6.3 Clause types and interpersonal management

Interrogatives, imperatives and declaratives behave differently in the workplace interaction. Imperatives can give a direct task, interrogatives can formulate requests and make directives less blunt, and declaratives can make an assertion of information or decision. A corpus-based approach can be used to quantify distributions of clause types as well as their co-occurrence with politeness markers and modality.

Again, multiword expressions may be central. For example, "can you" and "could you" are interrogative based on directive routines (this makes more sense for the analysis of phraseology and grammar together) (Kallens & Christiansen, 2022).

6.4 Studying Variation in the Pragmatic and Discourse-Level

Pragmatic variation deals with how utterances serve social activities not limited to the literal sense such as by

requesting, apologizing, aligning, disagreeing, and negotiating. Workplace pragmatics in particular is sensitive to power and normativity, because a lot of communication is done to coordinate action in a way that will not ruin relationships.

6.5 Mitigation, hedging, and stance

Mitigation and hedging can work to soften directives, communicate openness and minimize interpersonal friction. They are also engaged in the management of epistemic stance, particularly where the information offered is uncertain, risk exists or where decisions are contested. Register theory would predict that such features would differ according to genre and task: risk updates should have greater markers of uncertainty, while compliance directives should have fewer. (Pescuma et al., 2023).

Stance can also be indicated by multiword expressions which structure commitments and suggestions. The theoretical argument that multiword expressions lie at the heart of organization in language suggests that the analysis of stance should be undertaken in terms of recurrent patterns and not of isolated markers. (Kallens & Christiansen, 2022). For example, "it seems that," "I think" or "my understanding is" are stance routines which may be distributed differently on role levels and channels.

6.6 Politeness and relational work

Politeness markers, expressions of gratitude and affiliation acknowledgements have important functions in the coordination of the workplace. They can lower the friction, ensure cooperation and demonstrate respect. Their distribution may be influenced by both role expectations and norm policing since bottom-up prescriptivism may enforce what is acceptable as tonation (Lukač & Heyd, 2023).

Digitally mediated channels can reach beyond words for politeness and affiliation - through emoji, reaction icons and short acknowledgements. Bilaniuk 2023 explains how digital circulation can play a role in the creation of communities and solidarity.

Workplace settings: In these settings, there may also be similar context specific mechanisms at work when teams take on board some shorthand that confirms membership. Inclusion-oriented approaches tend to suggest that such



practices can lead to inclusion or exclusion, depending on who is at ease with or has access to the norms. (Wei, 2023).

6.7 Discourse organization and alignment

Alignment strategies such as repetition, confirmation tokens, shared framing, and consensus-oriented language such as, "align," "sync," "on the same page," "agree," etc. These can be used as tools for collaborative decision making and conflict reduction. From a functional communication perspective, alignment is part of interactive communication and therefore it should not be surprising that it should vary according to task and context. (Doedens & Meteyard, 2022). A register framework goes further, by predicting that alignment language is genre sensitive: one might find more cross functional coordination emails, and less policy documents which express, to use this example, parallel alignment. (Pescuma et al., 2023).

A pragmatic analysis may also be concerned with correction and the enforcement of norms. Grass-roots prescriptivism emphasises the role of peers to actively speak police language and this will potentially influence the way people frame their requests, corrections, or disagreements. (Lukač & Heyd, 2023). In the case of workplace corpora, this can be operationalized by coding for correction sequences, meta linguistic comments, and references to "tone" or "wording" and testing whether this is distributed in an uneven manner across gender and role.

7. GENDERED PATTERNS IN RELATION TO WORKPLACE POWER STRUCTURES

Interpreting patterns of language across genders without taking account of hierarchy runs the risk of perpetuating stereotypes. Register and functional view of communication soundly implies that language choices are determined by situational constraints and communicative objectives as well as hierarchical is a major situational variable in workplaces. (Doedens, M., & Meteyard, L. (2022). This section therefore specifies interpretive principles regarding observed gender linked patterns to power structures.

7.1 Hierarchy as a mediating variable

A fundamental and analytic principle is the principle of mediation: If differences in gender in the use of directives or obligation language vanish when the influence of role

and seniority is controlled, one may properly be looking at organizational power rather than gender. This does refract the methodological emphasis on distinguishing between person differences from situational variability. (Pescuma et al., 2023). Conversely, if the persistence of gender linked differences within role strata is the case, then again gender may operate through differentiated interactional expectations within the same hierarchical level, possibly through norm policing or differential evaluations of similar speech acts.

7.2 Normativity, evaluation, and linguistic authority

Workplace power is not just in a formalistic sense; it is also in an ideological sense. Expectations about "professional English": Possessing or lacking access to an insider position ("gatekeeping")? According to Hamid (2022) hegemonic native speakers or standard norms can limit linguistic legitimacy and decolonising knowledge requires trouble to these norms. (Hamid, 2022). By way of analogy, in workplace settings, ideologies of standardness can define what is possible and permissible as "clear," "confident," "professional," and potentially discipline communicative styles that may, rightly or wrongly, be gendered.

Grassroots prescriptivism further complicates the issue of authority because it demonstrates that language policing can be mobilized by peers, as well as the participatory culture putting into new dynamics on norm enforcement. (Lukač & Heyd, 2023). For example, in the workplace this might be in the form of colleagues correcting tone, punctuation or phrasing, or advising others to be "more direct" or "less harsh." These corrections are not neutral and work to encode ideologies of appropriate workplace discourse and can be especially burdensome for certain groups.

7.3 Channel and surveillance effects

Power is also influenced by the channel of communication. Emails and documents form durable records which can be audited, chat is fast but can be searched, nonetheless. Research on the ethics and data practices involved in language technologies points to the way that assumptions of data collection and data reuse can raise privacy and representations issues. (Reitmaier et al., 2022). In the context of communication at work, the awareness of permanence and potential for surveillance can



affect language choices, such as an increase in formality, avoidance of humour or more mitigation. These dynamics can be combined with gendered vulnerability to scrutiny, a possibility which should be tested empirically via channel comparisons and interaction models.

8. DISCUSSIONS

Given the limited direct gender focused corpus literature within the provided sources, the focus of integration is on how patterns observed at the workplace, when calculated, can be interpreted in terms of register, functional communication, ideology and phraseology frameworks.

Register theory offers mainly an interpretive point of reference. If, on the other hand, gender linked differences cluster in particular genres, such as chat and meeting talk, this would support the view that situational function is responsible for linguistic variability and that formal registers limit stylistic dispersion. (Pescuma et al., 2023). If differences are maintained in multiple genres, then it would be more appropriate to argue for a greater claim relating to pervasive gender linked routines, however one must be cautious in interpretation.

Functional communication theory does support taking linguistic features as resources for accomplishing situated goals. For instance, higher rates of mitigation may be indicative of the interactional strategies for coordination in interdependent teams rather than lack of confidence. (Doedens & Meteyard, 2022). Similarly, grammatical choices associated with modality and agency may be understood as a means of dealing with obligation, risk and responsibility in particular communicative situations.

The theory of phraseology and multiword expressions contributes to reinforcing the interpretation of routines that take place recurrently in the workplace. Because multiword expressions are argued to play a key role in linguistic knowledge and to bridge lexicon and grammar, gender linked differences in recurrent routines may be more meaningful than difference in single tokens. (Kallens & Christiansen, 2022). For example, gender related variation in request formulas/alignment phrases might hint at differences in the routine framing of organizational actions, which might be expressions of interactional norms and role distributions.

Ideology based perspectives uphold the critical interpretation about whether the observed patterns are a result of differential norm pressures. Hamid (2022) calls for

breaking from hegemonic norms of standard and this way of thinking promotes asking how ideal "professionalism" norms in the workplace are implicitly favor certain styles. (Hamid, 2022). Lukač and Heyd (2023) make a similar point about changing dynamics of linguistic policing that can inform how meta pragmatic commentary is to be interpreted in texts from the workplace, such as comments on tone or correctness. (Lukač & Heyd, 2023).

Finally, inclusion-oriented perspectives recommend that workplace communication findings be interpreted in the light of equity. Wei (2023) puts translanguaging as a part of overall inclusion and social justice agenda which calls for mindset change and not practice at the surface. (Wei, 2023). Applied to gender inclusive communication, this means that interventions should not simply prescribe how to rename things but also change the ideologies of the workplace that influence how language is judged and whose voices are taken as authoritative.

9. PRACTICAL AND POLICY IMPLICATIONS

Workplace language policies are often designed to create a clearer and more efficient language, although they also determine who is included and who belongs. The implications below are cast in terms of actionable recommendations based on the theoretical synthesis and on the type of empirical patterns corpus analysis is apt to reveal.

9.1 Evidence informed communication training

Communication training can be furnished by a corpus-based evidence base that distinguishes (a) genre appropriate norms from (b) avoidable prescriptive unreasonableness. As there is a risk that bottom-up prescriptivism will lead to peer policing of language, training should address the implications that "professional tone" judgments can be subjective and socially patterned in an explicit manner. (Lukač & Heyd, 2023). This can also take a functional framing when training may teach "be more direct" rather prescriptive as choosing directive strength that fits the task urgency and relationship context. (Doedens & Meteyard, 2022).



9.2 Inclusive language guidelines as ideological change

Inclusive communication guidelines should be considered as an ideological project instead of a checklist added on. Wei (2023) emphasizes that the process of inclusion-oriented practice needs a shift in mindset, and not just the superficial form of techniques. (Wei, 2023). In workplaces this implies guidelines which (a) promote reflection on the way language norms may privilege certain styles of communication, (b) offer examples of respectful alternatives to certain styles within different registers; and (c) acknowledge that it is possible to use multiple styles of communication, depending on the context.

Hamid (2022) emphasizes the relevance of pushing against hegemonic norms within the use of English, which bolsters the argument of not settling with a single standard of "professional English" as neutral. (Hamid, 2022). in international working environments, inclusive language policy should cater for different Englishes and discourage stigmatizing nonstandard forms of English which are intelligible, especially when such stigmatization intersects with gender, race, or other identities.

9.3 Ethical governance of workplace language data

Because communication corpora in the workplace deal with sensitive information, the governance of the organization should ensure that language data is not being repurposed for punitive surveillance without a measure of transparency and safeguards. Research on language technologies raises issues of data privacy and problematic assumptions about data interchangeability and data reuse. (Reitmaier et al., 2022). Even if the current research is anonymized analysis, attention needs to be paid to data access, data retention, data consent mechanisms, and acceptable uses of linguistic analytics in policies.

Technology oriented studies also put emphasis on the shared ethical responsibility and awareness for integrating AI tools. (Alhalangy & AbdAlgane, 2023). In workplaces, this means clear documentation of any automated text processing for analytics purposes and bias audits where appropriate and participatory consultation with employees about the methods of use of communication data.

10. ARTICULATING LIMITATIONS AND FUTURE RESEARCH DIRECTIONS

Several limitations arise from the study design and the limitations of the selection of the reference set.

- First, the operationalization of gender depends on metadata that is available. If gender is obtained in binary categories, nonbinary and fluid identities are not represented, which is a limitation to the interpretive validity. Future work needs to create ethically sound methods to both represent the diversity of gender and to analyse the interaction between gender identity and the languages used at work.
- Second, corpus representativeness is a permanent risk. Even with stratified sampling, texts that have been collected may oversample the occurrences of certain roles, departments, or types of communication. Broader computational scholarship helps us to be warned that datasets are unreliable in terms of representing how people communicate, and that practices of data collection can also be biased and privacy threatening. (Aji et al., 2022; Reitmaier et al., 2022). Future research should incorporate explicit audits of representativeness, dispersion measures as well as sensitivity analyses testing whether findings are similar for subsets.
- Third, the reference set provided has limited direct scholarship writing on gendered discourse in the workplace and this places limits on how strong ideological literature-based claims about expected gender patterns can be. Future work should embody directly pertinent gender and workplace discourse scholarship to better both theoretically groundwork and to contextualize findings more fully within established debates.
- Fourth, channel differences can be huge. Because written like chat can be like speech and because both workplace communication and research are increasingly mediated through technology, future research should expand the type of data collected and



thus include voice notes and hybrid meeting discourse while considering ethical principles. Adjacent work on speech and technology does serve to emphasize the influence of tools and situations on practices of communication. (Alhalangy & AbdAlgane, 2023; Reitmaier et al, 2022; Sun, 2023).

11. CONCLUSION

Workplace discourse is a consequential domain for the study of linguistic variation given that it simultaneously performs both a practical coordination and institutionalizes social relations. Treating communication in the workplace as functional and situated aids the analysis of language features as resources for accomplishing tasks and managing relationships. (Doedens & Meteyard, 2022). Register theory further predicts that linguistic variability is in turn conditioned by situation and function and thus motivates multi genre corpora along with complementary methods to separate individual variability from the effects of the situation. (Pescuma et al., 2023). Multiword expression theory enhances this type of approach by emphasizing that formulaic phraseology lies midway between lexicon and grammar and may encode conventionalized workplace routines where gendered connected differences may be more obvious than isolated tokens. (Kallens & Christiansen, 2022).

Within this theoretical frame, the best way to think about gendered language use in the workplace is as an empirical question one needs to carefully design a corpus with transparent metadata designed and modeled in ways that consider role and hierarchy. Normativity and ideology also need to be considered as workplace "professional English" can act as a gatekeeping standard as well as peer driven prescriptivism can contribute to the evaluation of tone and correctness. (Hamid, 2022; Lukač & Heyd, 2023). Finally, inclusion-oriented perspectives suggest that the value of such research will not center simply on describing differences but instead will revolve around informing equitable communication practices and policies for reducing unjust burdens and increasing legitimate ways of speaking and writing at work. (Wei, 2023).

REFERENCES

Aji, A. F., Winata, G. I., Koto, F., Cahyawijaya, S., Romadhony, A., Mahendra, R., Kurniawan, K.,

Moeljadi, D., Prasajo, R. E., Baldwin, T., Lau, J. H., & Ruder, S. (2022). One country, 700+ languages: NLP challenges for underrepresented languages and dialects in Indonesia. *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 7226 to 7249. <https://doi.org/10.18653/v1/2022.acl-long.500>

Alhalangy, A., & AbdAlgane, M. (2023). Exploring the impact of AI on the EFL context: A case study of Saudi universities. *Journal of Intercultural Communication*, 41 to 49. <https://doi.org/10.36923/jicc.v23i2.125>

Bilaniuk, L. (2023). Memes as antibodies. 143 to 166. <https://doi.org/10.4324/9781003382607-10>

Campoy, J. M. H., & Hernández Campoy, J. M. (2025). In Memoriam: William Labov. *Nexus*, 2025(1), 138 to 141. <https://doi.org/10.63182/nexus-2025-1.5.01>

Doedens, W., & Meteyard, L. (2022). What is functional communication? A theoretical framework for real world communication applied to aphasia rehabilitation. *Neuropsychology Review*, 32(4), 937 to 973. <https://doi.org/10.1007/s11065-021-09531-2>

Hamid, M. O. (2022). English as a Southern language. *Language in Society*, 52(3), 409 to 432. <https://doi.org/10.1017/s0047404522000069>

Kallens, P. C., & Christiansen, M. H. (2022). Models of language and multiword expressions. *Frontiers in Artificial Intelligence*, 5, 781962. <https://doi.org/10.3389/frai.2022.781962>

Lukač, M., & Heyd, T. (2023). Grassroots prescriptivism. 227 to 245. <https://doi.org/10.4324/9781003095125-16>

Pescuma, V. N., Serova, D., Lukassek, J., Sauermann, A., Schäfer, R., Adli, A., Bildhauer, F., Egg, M., Hülk, K., Ito, A., Jannedy, S., Kordoni, V., Kuehnast, M., Kutscher, S., Lange, R., Lehmann, N., Liu, M., Lütke, B., Maquate, K., Mooshammer, C., Morteza pour, V., Müller, S., Norde, M., Pankratz, E., Patarroyo, A. G., Pleşca, A., Ronderos, C. R., Rotter, S., Sauerland, U., Schnelle, G., Schulte, B., Schüppenhauer, G., Sell, B., Solt, S., Terada, M., Tsiapou, D., Verhoeven, E., Weirich, M., Wiese, H., Zaruba, K., Zeige, L. E., Lüdeling, A., & Knoeferle, P. (2023). Situating language register across the ages, languages, modalities, and cultural aspects: Evidence from complementary methods. *Frontiers in Psychology*, 13, 964658 to 964658. <https://doi.org/10.3389/fpsyg.2022.964658>



Reitmaier, T., Wallington, E., Raju, D. K., Klejch, O., Pearson, J., Jones, M., Bell, P., & Robinson, S. (2022). Opportunities and challenges of automatic speech recognition systems for low resource language speakers. *CHI Conference on Human Factors in Computing Systems*, 1 to 17. <https://doi.org/10.1145/3491102.3517639>

57(2), 203 to 214.
<https://doi.org/10.1017/s0261444823000186>

Wei, L. (2023). Transformative pedagogy for inclusion and social justice through translanguaging, co learning, and transpositioning. *Language Teaching*,